AU-System Radio has been an active participant in the WAP Forum standardization work since it started in June 1997. We offer highly skilled specialist consultants within the following WAP related areas: WAP business and service development; WAP infrastructure development; WAP service creation / development; and WAP client application development.

WAP
White Paper

February, 1999

...when time is of the essence

Contents

AU-SYSTEM

Executive Summary

Mobile networks of today do often not provide desired flexibility when value added services are about to be introduced. This since it often is a rather complicated and lengthy task to launch such services. The Wireless Application Protocol (WAP) addresses this issue by introducing the concept of the Internet as a wireless service platform.

The Internet has proven to be an easy and efficient way of delivering services to millions of "wired" users. In 1997, Ericsson, Motorola, Nokia, and Unwired Planet took the initiative to found WAP Forum, which mission is to bring the convenience of the Internet into the wireless community as well. WAP Forum has today gained great credence in the wireless industry all over the world, more than 90 of the world's leading companies in the business of wireless telecommunication are members as of February 1999.

By addressing the constraints of a wireless environment, and adapt existing Internet technology to meet these constraints, the WAP Forum has succeeded in developing a standard that scale across a wide range of wireless devices and networks. Key features offered by WAP are:

- **A programming model similar to the Internet's**
  Re-use of concepts found on the Internet enables a quick introduction of WAP based services since both service developers and manufacturers are familiar with these concepts today.
- **Wireless Markup Language (WML)**
  A markup language used for authoring services, fulfilling the same purpose as HyperText Markup Language (HTML) do on the World Wide Web (WWW). In contrast to HTML, WML is designed to fit small handheld devices.
- **WMLScript**
  WMLScript can be used to enhance the functionality of a service, just as for example JavaScript may be utilised in HTML. It makes it possible to add e.g. procedural logic and computational functions to WAP based services.
- **Wireless Telephony Application (WTA)**
  The WTA framework defines a set of features that provides a means to create telephony services. This is accomplished by introducing an in-client interface to the mobile network, handling of network events, a repository that allow real-time handling of services, and a mechanism supporting server initiated services.
- **Optimised protocol stack**
  The protocols used in WAP are based on well-known Internet protocols such as HyperText Transport Protocol (HTTP) and Transport Control Protocol (TCP), but have been optimised to address the constraints of a wireless environment, such as, low bandwidth and high latency.

The opportunity of creating wireless services on a global basis will attract operators as well as third party service providers, resulting in both co-operations and competition that do not exist today. WAP provides a means to create not only services that we are used to from the World Wide Web today, but also telephony services.

Introduction

The Wireless Application Protocol (WAP) is a result of joint efforts taken by companies teaming up in an industry group called WAP Forum. The objective of the forum is to create a license-free standard that brings information and telephony services to wireless devices. To access these services WAP utilise the Internet and the World Wide Web (WWW) paradigm.

WAP scales across a broad range of wireless networks, implying that it has the potential to become a global standard and that economies of scale thus can be achieved.

In order to provide wireless access to the information space offered by the WWW, WAP is based on well-known Internet technology that has been optimised to meet the constraints of a wireless environment.

Background

WAP could roughly be described as a set of protocols that has inherited its characteristics and functionality from Internet standards and standards for wireless services developed by some of the world's leading companies in the business of wireless telecommunications.

In 1995 Ericsson initiated a project which purpose was to develop a general protocol, or rather a concept, for value added services on mobile networks. The protocol was named Intelligent Terminal Transfer Protocol (ITTP), and handles the communication between a service node where the service application is implemented and an intelligent mobile telephone. The ambition was to make ITTP a standard for value added services in mobile networks.

During 1996 and 1997 Unwired Planet, Nokia, and others, launched additional concepts in the area of value added services on mobile networks.

Unwired Planet presented Handheld Device Markup Language (HDML) and Handheld Device Transport Protocol (HDTP). Just as HyperText Markup Language (HTML) used on the WWW, HDML is used for describing content and user interface, but optimised for wireless Internet access from handheld devices with small displays and limited input facilities. In the same manner HDTP could be considered to be a wireless equivalent of the standard Internet HyperText Transport Protocol (HTTP), i.e. a lightweight protocol to perform client/server transactions.

In March 1997 Nokia officially presented the Smart Messaging concept, an Internet access service technology specially designed for handheld GSM devices. The communication between the mobile user and the server containing Internet information uses Short Message Service (SMS) and a markup language called Tagged Text Markup Language (TTML). Just like HDML, this language is adapted for wireless communication, i.e. narrowband connections.

With a multitude of concepts there was a substantial risk that the market could become fragmented, a development that neither of the involved companies would benefit from. Therefore, the companies agreed upon bringing forth a joint solution. WAP was born...

WAP Forum

On June 26 1997 Ericsson, Motorola, Nokia, and Unwired Planet took the initiative to start a rapid creation of a standard for making advanced services within the wireless domain a reality. In December 1997 WAP Forum was formally created, and after the release of the WAP 1.0 specifications in April 1998, WAP Forum membership was opened to all. Today (February 1999) over 90 companies are members of WAP Forum. Among these companies many of the world's leading terminal and infrastructure manufacturers, software companies, operators, and service providers are found. The handset manufacturers in WAP Forum represent over 90% of the world market across all technologies, and the network operators are representing about 100 million subscribers.

The main objectives of the WAP Forum are:

- Independent of wireless network standard
- Open to all
- Will be proposed to the appropriate standards bodies
- Applications scale across transport options (GSM, IS-95, IS-136, PDC, etc)
- Applications scale across device types (mobile phones, PDAs, etc)
- Extensible over time to new networks and transports (e.g. 3G systems)

WAP- The Wireless Service Enabler

The following sections discuss why the Internet is a suitable platform for wireless value added services and the tight coupling between the Internet and WAP programming models.

Why Using the Internet?

During the last couple of years we have become used to the wide variety of services offered by the Internet and the WWW. Not only the services themselves attract us, but also the convenient way of accessing them via an Internet browser. We can access the same services all over the world as long as we have access to a computer and the Internet. Nobody can deny that the Internet has scored a tremendous success, nearly a 150 million users (end of 1998, source: Computer Industry Almanac) can not be wrong.

Service providers do also benefit from the WWW paradigm since their services can be deployed independent of the location of the users. The services are created and stored on a server, meaning that it becomes very easy to change them according to the needs of the customers. By using off-the-shelf authoring-tools services are created with minimum effort which, combined with the fast and convenient way of launching them, enables an extremely short time-to-market. The reduced service development time does of course also imply reduced costs compared to conventional service development in wireless networks.

As users become more and more dependent of services offered on the Internet, one shortcoming becomes increasingly evident - the need for a wire to connect to the Internet. This shortcoming make itself especially remembered to the millions of users that spend a substantial amount of their time on the move.

The last few years attempts to make this shortcoming disappear has not made the Internet crossing the chasm as a wireless service platform, only the early adopters has accepted the technologies provided so far. One of the main reasons for this is the lack of a widely accepted standard, a problem being addressed by the WAP Forum. The wide support for WAP will most probably enable the Internet as a means to provide services to wireless devices within a foreseeable future. This includes both services that we recognise from the WWW, and services like telephony services. The two next sections show how the Internet and WAP relate to each other.

The Internet Model

The Internet model makes it possible for a client to reach services on a large number of origin servers; each addressed by a unique Uniform Resource Locator (URL). The content stored on the servers is of various formats, but HTML is the predominant. HTML provides the content developer with a means to describe the appearance of a service in a flat document structure; i.e. the entire content of a page is shown simultaneously. If more advanced features like procedural logic are needed, scripting languages such as JavaScript or VB Script may be utilised.

The figure below shows how a WWW client request a resource stored on a web server. As mentioned above, a resource on the Internet is identified by a unique URL, that is, a text string constituting an address to that resource. In the example below, the resource is an HTML document.

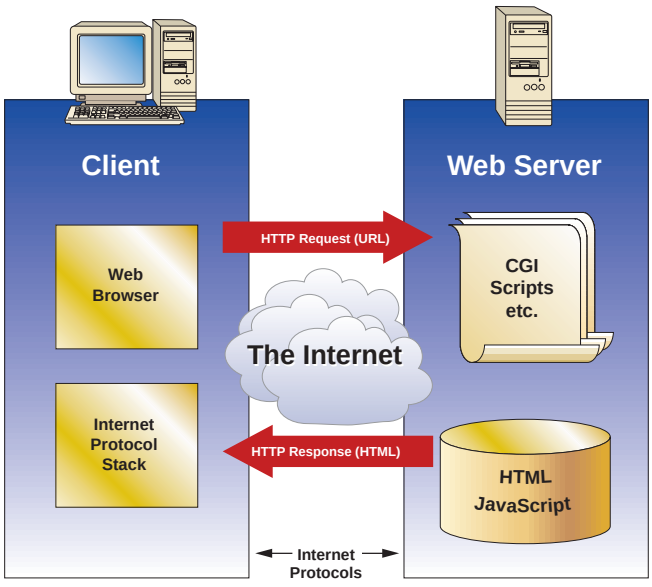

Figure 1 : The Internet Model

On the Internet, standard communication protocols, like HTTP and Transmission Control Protocol/Internet Protocol (TCP/IP) are used.

The content may be static or dynamic. Static content is produced once and not changed or updated very often, for example a company presentation. Dynamic content is needed when the information provided by the service changes more often, for example timetables, news, stock quotes and account information. Technologies such as Active Server Pages (ASP), Common Gateway Interface (CGI), and Servlets allow content to be generated dynamically.

The WAP Model

WAP do also make use of the Internet paradigm to provide a flexible service platform. In order to accommodate wireless access to the information space offered by the WWW, WAP is based on well-known Internet technology that has been optimised to meet the constraints of a wireless environment.

Services created using HTML would not fit very well on small handheld devices since they are intended for use on desktop computers with big screens. Low bandwidth wireless bearers would neither be suitable for delivering the rather extensive information that HTML pages often consist of. Therefore a markup language adapted to these constraints has been developed - the Wireless Markup Language (WML).

WML offers a navigation model designed for devices with small displays and limited input facilities (no mouse and limited keyboard). In order to save valuable bandwidth in the wireless network, WML can be encoded into a compact binary format. Encoding WML is one of the tasks performed by the WAP Gateway/Proxy, which is the entity that connects the wireless domain with the Internet.

WAP do also provide a means for supporting more advanced tasks, comparable to those solved by using for example JavaScript in HTML. The solution in WAP is called WML Script.

The figure below shows the WAP programming model. Note the similarities with the Internet model. Without the WAP Gateway/Proxy the two models would have been practically identical.

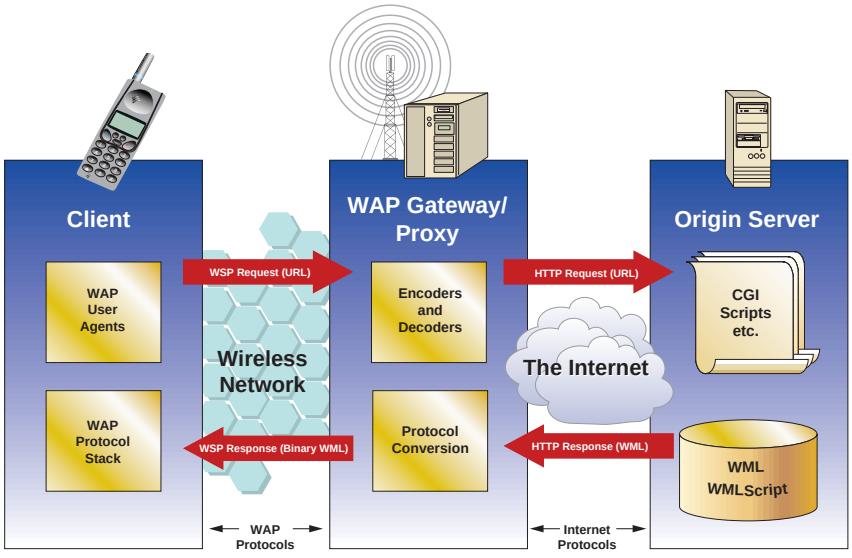

Figure 2 : The WAP Model

The observant reader may note that the request that is sent from the wireless client to the WAP Gateway/Proxy uses the Wireless Session Protocol (WSP). In its essence, WSP is a binary version of HTTP.

WAP is designed to scale across a broad range of wireless networks, like GSM, IS-95, IS-136 and PDC. Finally, the WAP protocol stack is designed in a layered fashion, meaning that it becomes extensible and future proof.

WAP Architecture

This chapter will give a brief overview of the WAP architecture, and the two following chapters will introduce the WAP application environment and the WAP protocols respectively. The descriptions are by no means complete; they rather outline the most important features provided by WAP. For a thorough description, please see the WAP specification suite available at http://www.wapforum.org.

WAP is designed in a layered fashion in order to be extensible, flexible, and scalable. With the Open System Interconnection model (OSI model) in mind, the WAP-stack basically is divided into five layers. They are:

- **Application Layer**
  Wireless Application Environment (WAE)
- **Session Layer**
  Wireless Session Protocol (WSP)
- **Transaction Layer**
  Wireless Transaction Protocol (WTP)
- **Security Layer**
  Wireless Transport Layer Security (WTLS)
- **Transport Layer**
  Wireless Datagram Protocol (WDP)

Each layer of the WAP protocol stack specifies a well-defined interface to the layer above, meaning that a certain layer makes lower layers invisible to the layer above. The layered architecture allows other applications and services to utilise the features provided by the WAP-stack as well. This makes it possible to use the WAP-stack for services and applications that currently are not specified by WAP.

The figure below shows the WAP protocol stack and how it relates to the protocols on the Internet.

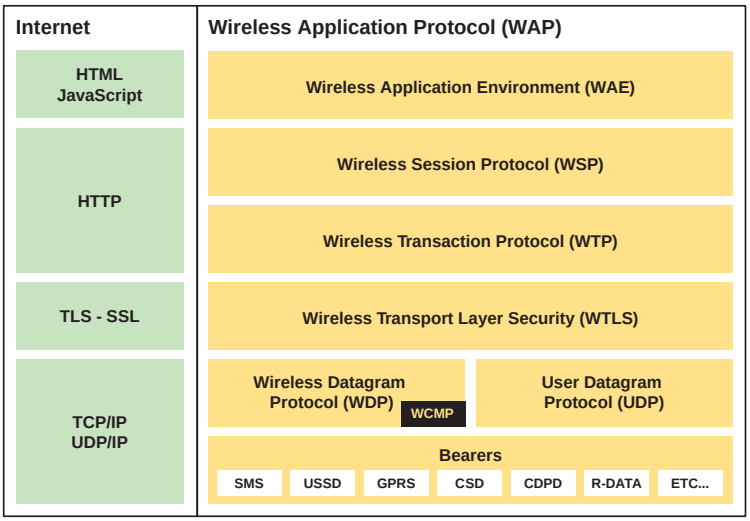

Figure 3 : The WAP Architecture

Note that the mobile network bearers in the lower part of the figure above are not part of the WAP protocol stack.

WAP Application Environment

The uppermost layer in the WAP stack, the Wireless Application Environment (WAE) provides an environment that enables a wide range of applications to be used on wireless devices. In the chapter "WAP - the wireless service enabler" the WAP WAE programming model was introduced. This chapter will focus on the various components of WAE:

- **Addressing model**
  A syntax suitable for naming resources stored on servers
- **Wireless Markup Language (WML)**
  A lightweight markup language designed to meet the constraints of a wireless environment with low bandwidth and small handheld devices
- **WMLScript**
  A lightweight scripting language
- **Wireless Telephony Application (WTA, WTAI)**
  A framework and programming interface for telephony services

In order to make use of the features mentioned above, WAP assumes that two user-agents will be available in the wireless device.

A user-agent is within this context typically an in-device application that interprets content in a well-defined manner and handles user-interactions when necessary. In the wired world, a user-agent is typically referred to as a browser; e.g. Microsoft Internet Explorer and Netscape Navigator used on desktop computers. As mentioned, WAP assumes two user-agents, the WML user-agent and the WTA user-agent.

The wording "WML user-agent" might be a little confusing since this user-agent not only interprets WML, but also WMLScript etc. The wording should rather reflect that this user-agent, in contrast to the WTA user-agent, is not capable of handling telephony services.

Addressing Model

WAP use the same addressing model as the one used on the Internet, that is, Uniform Resource Locators (URL). A URL uniquely identifies a resource, e.g. a WML document, on a server that can be retrieved using well-known protocols.

In addition to URLs, WAP also uses Uniform Resource Identifiers (URI). A URI is used for addressing resources that are not necessarily accessed using well-known protocols. An example of using a URI is local access to a wireless device's telephony functions.

Wireless Markup Language (WML)

The Wireless Markup Language is WAP's analogy to HTML used on the WWW. WML is based on the Extensible Markup Language (XML).

WML use a deck/card metaphor to specify a service. A card is typically a unit of interaction with the user, that is, either presentation of information or request for information from the user. A collection of cards is called a deck, which usually constitutes a service. This approach ensures that a suitable amount of information is displayed to the user simultaneously since inter-page navigation can be avoided to the fullest possible extent.

Key features of WML include:

- Variables
- Text formatting features
- Support for images
- Support for soft-buttons
- Navigation control
- Control of browser history
- Support for event handling (for e.g. telephony services)
- Different types of user interactions, e.g. selection lists and input fields

WML can be binary encoded by the WAP Gateway/Proxy in order to save bandwidth in the wireless domain.

WMLScript

WMLScript is based on ECMAScript, the same scripting language that JavaScript is based on. It can be used for enhancing services written in WML in the way that it to some extent adds intelligence to the services, for example procedural logic, loops, conditional expressions, and computational functions.

WMLScript can be used for e.g. validation of user input. Since WML does not provide any mechanisms for achieving this, a round-trip to the server would be needed in order to determine if user input is valid or not if scripting was not available. Access to local functions in a wireless device is another area where WMLScript is used; for example access to telephony related functions.

WMLScript do also support WMLScript Libraries. These libraries contain functions that extend the basic WMLScript functionality. This provides a means for future expansion of functions without having to change the core of WMLScript.

Just as with WML, WMLScript can be binary encoded by the WAP Gateway/Proxy in order to minimise the amount of data sent over the air.

Wireless Telephony Applications (WTA)

The Wireless Telephony Application (WTA) environment provides a means to create telephony services using WAP. As already mentioned, WTA utilise a user-agent separate from the common WML user-agent, at least logically. The WTA user-agent is based on the WML user-agent, but is extended with functionality that meets the special requirements for telephony services. This functionality include:

- **Wireless Telephony Application Interface (WTAI)**
  An interface towards a set of telephony-related functions in a mobile phone that can be invoked from WML and/or WMLScript. These functions include for example: call-management, handling of text messages and phonebook control.
  WTAI is divided into three categories: Network Common Functions, Network Specific Functions, and Public Functions. The common functions are available in all types of networks, while the specific functions specify functions that are unique to a certain network type. In contrast to the other two function libraries, the Public Functions library can be invoked from the WML user-agent as well. Currently, the Public Functions library only contains a function for setting up calls, which must be acknowledged by the user before it is carried out.
- **Repository**
  Many WTA services put requirements on real-time handling, implying that it is not feasible to retrieve content from a server since this involves a certain delay. The repository makes it possible to store WTA services persistently in the device in order to enable access to them without accessing the network.
- **Event handling**
  Typical events in a mobile network are incoming call, call disconnect, and call answered. In order to create telephony services, it must be possible to handle these events. The event handling within WTA enables WTA services stored in the repository to be started in response to such events. Events can also be bound to a certain action in WML in order to make it possible to handle events within a service.
- **WTA Service Indication**
  A content type that allows the user to be notified about events of different kinds (e.g. new voice mails) and be given the possibility to start the appropriate service to handle the event. In its most basic form, the WTA Service Indication makes it possible to send a URL and a message to a wireless device. The message is displayed to the user, and she is asked whether she wants to start the service indicated by the URL immediately or if she wants to postpone the Service Indication for later handling. The WTA Service Indication should be delivered to a device using push, an area where WAP Forum currently is working actively.

WTAI enables access to functions that are not suitable for allowing common access to them (except for the Public Function Library). For instance, setting up calls and manipulating the phonebook without user-acknowledgement can both imply undesired costs and violate user-integrity if the corresponding functions are used inappropriately. The other functions provided by the WTA framework can be considered in the same way.

The WTA framework relies on a dedicated WTA user-agent capable of carrying out these functions, a functionality not provided by the common WML user-agent. Only “trusted” content providers should be able to make content available to the WTA user-agent, i.e. the operator or content providers trusted by the operator. Thus it must be possible to distinguish between servers that are allowed to supply the user-agent with services containing these functions, and those who are not. To accomplish this the WTA user-agent retrieves its services from the WTA domain, which, in contrast to the Internet, is controlled by the network operator. The figure below shows how WTA services and other services are separated from each other using WTA access control based on port numbers:

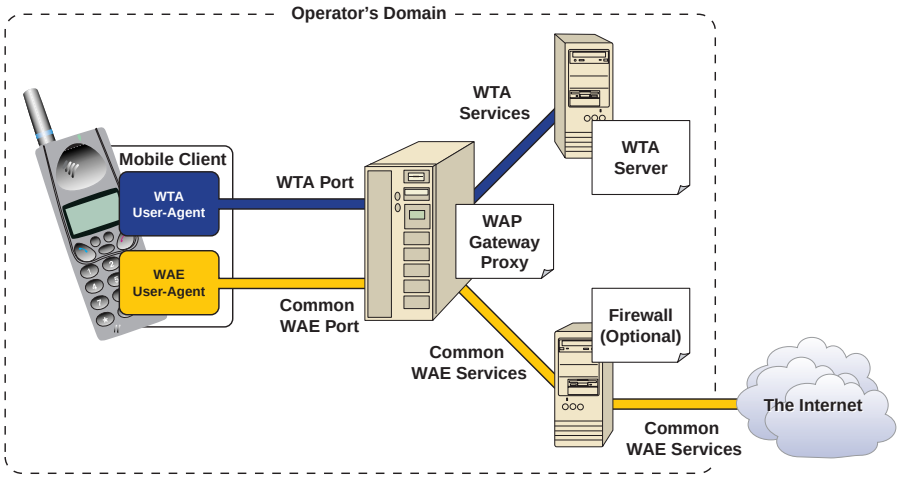

Figure 4 : WTA Access Control

The WTA server may be an ordinary web server used for housing content. It may also be able to communicate with other entities, such as IN-nodes or voice mail systems, to provide extended telephony-related functionality. The communication with such entities can be controlled by applications on the WTA server that can be referenced within a WTA service using URLs.

WAP Protocols

The WAP protocol suite contains four protocols for handling the communication between clients and the WAP Gateway/Proxy. These protocols are modelled after protocols used on the Internet, and can be used in four different configurations (each of the protocols mentioned below are explained in the next sections):

- **Connectionless mode**
  This configuration utilises only WSP on top of WDP. It offers a datagram service, meaning that sent messages are not acknowledged, and hence no guarantee of delivery is offered. The functionality can be seen as a simple "send-and-forget" model.
- **Connectionless mode with security**
  In addition to what is mentioned above, WTLS is used in order to provide authentication, encryption, etc.
- **Connection mode**
  The connection mode uses WTP in addition to WSP and WDP. WTP offers reliable transmissions, meaning that sent messages must be acknowledged and may be retransmitted if lost. It also uses a mode of WSP that handles long-lived sessions
- **Connection mode with security**
  In addition to what is mentioned above, WTLS is used in order to provide authentication, encryption, etc.

Wireless Session Protocol (WSP)

WSP is the interface between WAE and the rest of the protocol stack. WSP is a binary version of HTTP 1.1 with additions such as:

- Capability negotiation.
- Header caching
- Long-lived sessions
- Push

The two main stack configurations (connection and connectionless mode) are named after the session services offered by WSP. The connectionless session service is basically a thin layer that WAE can use when there is no need for reliable delivery of messages.

The main functionality of the connection mode of WSP is to set up a session between a client and the WAP Gateway/Proxy. This session handles capability negotiation at session establishment and also communication interrupts such as change of bearer. It is assumed to be long-lived and can be suspended, and later resumed, instead of disconnected if no communication will be needed for a time. This means that no new capability negotiation is needed when the session is resumed, which lessens the traffic load. Support for asynchronous handling of requests implies that, if several data requests have been sent, the answers may be delivered in a different order.

WSP also supports header caching in order to minimise bearer utilisation. In HTTP, which does not support header caching, about 90% of the requests contain static headers that need to be sent over and over again.

Wireless Transaction Protocol (WTP)

WTP is responsible for control of transmitted and received messages. It provides a reliable communication where messages are uniquely identified so as not to be accepted twice and may be retransmitted to the peer if lost in transmission. There is no connection between communications as every communication sequence is only alive during the exchange of an individual message set. WTP works with three different message classes.

- **Unreliable "send" with no result message**
  No retransmission if the sent message is lost.
- **Reliable "send" with no result message**
  The recipient acknowledges the sent message. Otherwise the message is resent.
- **Reliable "send" with reliable result message (three-way communication)**
  A data request is sent and a result is received which finally is acknowledged by the initiating part.

WTP is also adapted to the constraints of wireless bearers in that it minimises the protocol overhead by introducing functionality to minimise the number of (re-) transmissions, for example, message concatenation and acknowledgement of received data requests. WTP can be extended with functionality for segmenting and reassembling messages. This includes selective retransmission of lost segments.

WSP and WTP are modelled together in such a way that the different functions in WSP have a defined and consistent usage of the message classes in WTP.

Wireless Transport Layer Security (WTLS)

As its name implies, the purpose of WTLS is to provide *transport* layer security between a WAP client and the WAP Gateway/Proxy. WTLS is based on Transport Layer Security (TLS) 1.0 but optimised for narrowband communication channels. Key features include:

- Integrity through the use of Message Authentication Codes (MAC)
- Confidentiality through the use of encryption
- Authentication and nonrepudiation of server and client, using digital certificates

These features make it possible to certify that the sent data have not been manipulated by a third party, that privacy is guaranteed, that an author of a message can be identified, and that both parties can not falsely deny having sent their messages. A secure connection is set up with an establishment phase where negotiation such as parameter settings, key exchange and authentication is performed. Both parties can abort the secure connection during establishment or at any time later.

WTLS is optional and can be used with both the connectionless and the connection mode WAP stack configuration. If used, it is always placed on top of WDP.

Wireless Datagram Protocol (WDP)

The base of the WAP protocol stack is a datagram layer, WDP, offering a consistent interface to the upper layers of the stack.

If WAP is used over a bearer supporting User Datagram Protocol (UDP), the WDP layer is not needed. On other bearers, such as GSM SMS, the datagram functionality is provided by WDP. This means that whether WAP uses UDP or WDP, it is given a datagram service, which hide the characteristics of different bearers and provides port number functionality. If necessary WDP can also be extended with functionality for segmenting and reassembling datagrams that are to big for the underlying bearer.

It is also possible to extend WDP with an optional protocol for error reporting called Wireless Control Message Protocol (WCMP). This protocol can be used when WAP is not used on an IP bearer (IP has its own control message protocol). WCMP can also be used for informational and diagnostic purposes.

Motivations for WAP

During recent years, both the Internet and wireless voice communication have undergone wide and rapid acceptance. The unification of these two technologies, the wireless Internet, has however not enjoyed the same development even though the Internet provides a means for rapid service development, short time-to-market, ease-of-use, and convenient manageability.

Why a new Standard?

A legitimate question is why the use of wireless data capabilities, in this context wireless Internet access, has not followed the trends of neither the wireless voice communication nor the Internet.

Just as with many other things in life, expectations matter very much. Anyone who have tried to access the Internet by using a laptop and a cellular phone, knows that the expectations we have created by using the Internet at the office or at home are not fulfilled; as a matter of fact it is usually a quite tiresome experience.

WAP addresses this issue by being designed to meet the constraints of a wireless environment. Both limitations in the network and in the client are taken into consideration.

The following sections outlines some of the motivations behind WAP, often with parallels to existing Internet technology.

Adapting to the Bounds of the Wireless Network

WAP scales across a broad range of wireless networks and bearers. Hence it is designed to allow access to services via the Internet using simple SMS as well as fast packet-data networks such as General Packet Radio Service (GPRS). The most important issues in the network addressed by WAP are summarised below:

☑ **Low Bandwidth:** The problem with poor performance when using wireless bearers with low bandwidth becomes especially valid if the user is not well aware of what services to access. This since the service must not consume much bandwidth if it should be suitable for wireless access. The larger portion of the mass market, as well as many advanced users, are not aware of this. And besides, the users should simply not have to care about how the services they access are designed in order to have their expectations fulfilled.

WAP addresses this issue by minimising the traffic over the air-interface. WML and WMLScript are binary encoded into a compact form when sent over the air in order to minimise the number of bits and bytes.

WSP, WAP's equivalent to HTTP on the Internet, is also binary for the same reason. Moreover, it supports long-lived sessions, that can be suspended and resumed, and header caching; saving valuable bandwidth since session establishment can then be done rather seldom.

The Wireless Transaction Protocol WTP, the analogy to the Internet's TCP, is not only designed to minimise the amount of data in each transaction, but also the number of transactions.

☑ **High Latency:** Wireless networks have high latency compared to wired networks. This constraint is relevant in all of today's wireless networks, even for those providing high bandwidth. This is addressed in WAP by minimising the roundtrips between the wireless device and the wireless network. An asynchronous request/response model is also used.

In wired networks the low latency implies that requests and responses can be handled synchronously since the time between them most often do not affect the user experience. In wireless networks with high latency this is not a feasible approach, especially when using high latency bearers like SMS. This issue is among other features addressed by WSP by allowing requests and responses to be handled asynchronously, that is, a new request can be sent before the response to an earlier request has been received.

The application environment in WAP uses the concept of scripting, meaning that roundtrips between a client and a server can be avoided when it comes to e.g. validation of user input. The Wireless Telephony Application environment addresses latency (and low bandwidth) by introducing the repository, which is a persistent storage container used for housing services that should be started in response to a event in the mobile network (e.g. an incoming call). Since these services are available immediately, no roundtrips to the WTA server are needed and thereby real-time handling is made possible.

Compared to TCP, WTP needs a smaller number of transactions for each method invoked, i.e. keeps the number of roundtrips down.

☑ **Less Connection Stability/Unpredictable Bearer Availability:** Wired network access provides a more or less reliable connection to the network. That is not the case in wireless networks where bearers might be inaccessible for shorter or longer periods of time due to fading, lost radio coverage, or deficient capacity.

As already mentioned, the sessions supported by WSP are assumed to be long-lived. The problem mentioned above is in WSP addressed by allowing lost sessions to be resumed, even when dynamically assigned IP addresses are used.

The transaction layer in WAP, WTP, has been kept simple compared to TCP used on the wired Internet. Since no connection is set up the effects of lost bearer and times of inactivity are minimised.

The nature of a wireless connection implies that small segments of a message often are lost. WTP support selective retransmission of data, meaning that the only the lost segments are retransmitted, not the entire message as in TCP.

Adapting to the Bounds of the Wireless Device

WAP is targeted at handheld devices of various kinds. Services should be accessible from a Handheld PC as well as from a small phone. WAP addresses this fact by taking the following issues into consideration:

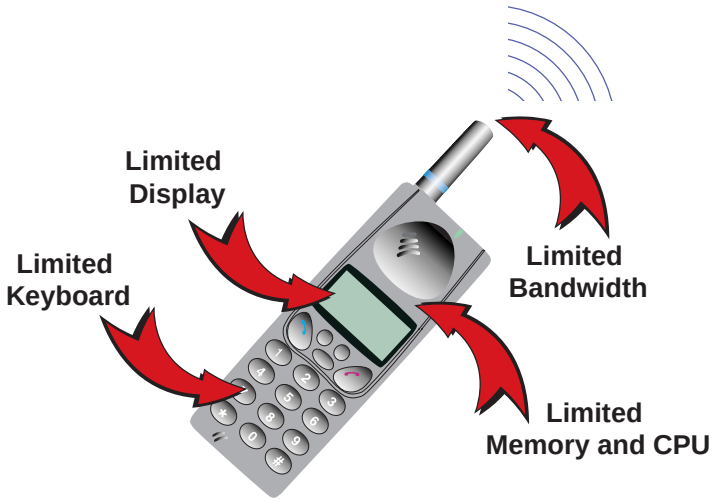

Figure 5 : Limitations of a Wireless Device

☑ **Small Display:** When accessing a service from a desktop computer, the size of the screen does not limit the user experience. Wireless devices might also have "big" displays, for example a Personal Digital Assistant (PDA). But many devices will have smaller displays, for example mobile phones, to provide larger portability. No matter how good these displays will be in the future, the size of the human hand will always limit the size of them. Try to imagine what the experience would be like when accessing a service designed for a device with a big display on a small phone. The result would in most cases be very bad performing services far from what was originally intended. The information the user really wants would probably be drowned in undesired information due to the low perspicuity that the small display of such a device offers.

Instead of using the flat document structure HTML provides, WML structures its document in decks and cards. A card is a single unit of interaction with the end-user, for instance a text-screen, a selection list, an input field, or a combination of them. A card is typically small enough to be displayed even on a small screen. When a service is executed the user navigates through a series of cards. The series of cards used for making a service is collected in a deck.

☑ **Limited Input Facilities:** Wireless devices do most often not have the same input facilities as their wired equivalents, that is, they lack QWERTY keyboards and have mouse-less interfaces.

WML addresses this issue as well. The elements that are used in WML can easily be implemented so they make very humble requirements on the keyboard. The use of decks and cards provides a navigation model that call for minimum inter-page navigation since the user is guided trough a series of cards instead of having to scroll up and down on a large page.

Further, soft-buttons are supported by WML in order to provide the service developer with a means to couple desired actions to vendor specific keys.

☑ **Limited Memory and CPU:** Wireless devices are usually not equipped with amounts of memory and computational power (CPU) comparable to desktop computers. The memory restriction is valid for RAM as well as for ROM. Even though the trend indicates that more memory and more powerful CPUs will be available in a foreseeable future, the relative difference will most probably remain.

WAP addresses these restrictions by defining a lightweight protocol stack adapted to its purpose. The limited set of functionality provided by WML and WMLScript makes it possible to implement browsers that make small claims on computational power and ROM resources. When it comes to RAM, the binary encoding of WML and WMLScript helps keeping the RAM as small as possible.

☑ **Limited Battery Power:** The stumbling block in wireless communication devices today is the operating time, i.e. the battery power restricts its usage. Even though batteries become better and better, and the radio interfaces are tuned to consume less power, there is still a lot to accomplish in this area.

Access to wireless services will increase the utilisation of bearers (radio interface), and thus will the power consumption also increase. This issue is solved by minimising the bandwidth (see above) and thus keeping the bearer utilisation as low as possible.

Roles and Benefits

The operator that runs a mobile network traditionally controls almost the entire value chain for mobile Value Added Services (VAS). Third party alliances are, if existing at all, in most cases restricted to providing the operator with information that it can use as a basis for serving its customers with mobile VAS. Typically, this information can be weather information from the national meteorological office, or stock quote information from a stockbroker. The network operator then conveys this information to the user, either automatically or when the users make an explicit request for it.

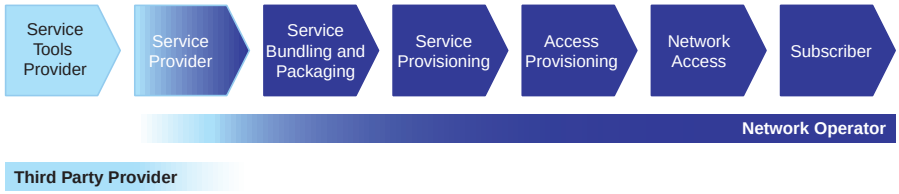

Figure 6 : Mobile VAS Value Chain Today

This scenario will most probably change when WAP enters the scene. Using the Internet as a service platform opens new possibilities for third party service providers to take part of the value chain at different stages. Third party service providers will be able to create WAP services, put them on the Internet, and thereby making them available to millions of subscribers. They will even be able to create complete suites of services and thus also affect the operator's role in bundling services. In co-operation with the operator, third party providers can seize the opportunity to take part of service and access provisioning as well. However, this requires that they make certain investments in network technology.

The operators must make important decisions about their role in providing mobile VAS when WAP is rolled out since the role of third party service providers is about to change rather dramatically. With the magnitude of new services that WAP will make available, network operators are unlikely to be able to serve all of its customers with self-made services that attracts each and every one of them. Therefor they have to decide how they should position themselves in the value chain, in order to be able to differentiate themselves from their competitors and have flexibility enough to respond to new preferences among its customers and changes on the market for mobile VAS in general.

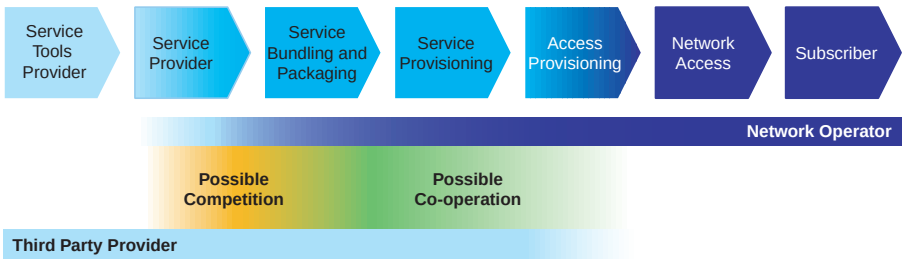

Figure 7 : Mobile VAS Value Chain with WAP

No matter to what degree the operator decides to co-operate with third party providers, it will still enjoy an increase in network utilisation, which of course will increase the revenues. Third party co-operation should be considered in order to maximise that utilisation and to provide an well-adapted mix of services that

allows the operator to differentiate itself from its competitors and attract new or underdeveloped market segments. This will likely reduce churn and improve customer loyalty, and thereby enable increased average revenue per user.

As indicated in the figure above, there is a possible risk of competition between the operator and third party providers when it comes to providing services. This since there is no hindrance for either of them to provide similar services. However, when services like telephony services are considered, i.e. services that require some level of integration with the wireless network, WAP provides the operator with a means to control how these services can be deployed.

The following sections outline how various groups may gain from WAP.

Subscribers

It is crucial that the subscribers will benefit from using WAP based services, otherwise there will be no incentive neither for WAP as a whole nor for any of the other groups mentioned below. The key-benefits can be summarised as:

- Portability
- Easy to use
- Access to a wide variety of services on a competitive market
- The possibility of having personalised services
- Fast, convenient, and efficient access to services
- To fulfil as many customers needs as possible, WAP devices will be available in various form factors, e.g. pagers, handheld PCs, and phones

Operators

As described in the introduction to this chapter, the network operator can cover the value chain to various degrees. Therefore, many of the advantages mentioned under “Service Providers” are be applicable to operators as well. The operator’s benefits may include:

- Address new market segments of mobile users by enabling a wider range of mobile VAS
- Deploy telephony services that in contrast to traditional telephony services are easy to create, update, and personalise
- Use the flexibility of WAP as a tool to differentiate from competitors
- Attractive interface to services will increase usage
- Increased revenues per user due to higher network utilisation
- Convenient service creation and maintenance, including short time-to-market
- Replace expensive customer care centres with WAP based services (E-care)
- WAP services are designed to be independent of the network, implying that an operator who runs different types of networks only have to develop its services ones
- An open standard means that equipment will be provided by many manufacturers

Service Providers

WAP opens new possibilities for service and content providers since they not necessarily have to come to an agreement with a specific operator about providing services to their customers. The gains are for example:

- Create a service once, make it accessible on a broad range of wireless networks
- Address new market segments by launching innovative mobile VAS. Keep old customers by adapting existing Internet services to WAP
- Keep old customers by adapting existing Internet services to WAP
- Convenient service creation and maintenance
- Creating a WAP service is no harder than creating an Internet service today since WML and WMLScript are based on well-known Internet technology
- Use standard tools like ASP or CGI to generate content dynamically
- Utilise existing investments in databases etc that are the basis of existing Internet services

Manufacturers

Mobile devices supporting WAP will be available in many different form factors, e.g. cellular phones, pagers, and handheld PCs. Hardware manufacturers will also need to supply operators etc with equipment, such as WAP Gateway/Proxys and WTA servers. According to WAP Forum, there will be 10's of millions of WAP enabled devices by the end of year 2000. Manufacturer benefits are for example:

- WAP scales across a broad range of mobile networks, meaning that WAP implementations can be used in devices supporting different types of networks
- The expected wide adoption of WAP implies that economies of scales can be achieved, meaning that the huge mass-market can be addressed
- The fact that WAP is designed to consume minimal amount of memory, and that the use of proxy technology relieves the CPU, means that inexpensive components can be used in the handsets
- Reuse the deep knowledge about wireless network infrastructure to develop advanced servers that seamlessly integrates mobile VAS with telephony
- Seize the opportunity to introduce new innovative products

Tools Providers

Today there is a large amount of tools available for creating applications for the web. Content developers have become used to the convenience that tools like FrontPage and DreamWeaver provides. The knowledge of how to develop these tools may be leveraged for developing tools supporting WAP as well. Tools providers will be able to:

- Reuse and modify existing products to support WAP, or even integrate WAP support in existing tools
- Address a new customer base in the wireless community

Services using WAP

When using the Internet from a desktop computer, it is very easy to find new and hopefully interesting services by using search engines, clicking on links and banners, typing in URLs recommended by a friend, etc. This is primarily a fact due to the attractive user-interface (a big screen, a full-sized keyboard, and a mouse). We simply sit down and "surf the net"...

With WAP it is different. While we are on the move, we do not want to go and look for the services we want. We just want the utility they provide, wherever we are. This requires a mindset different to what we are used to from the WWW today. Instead of using advanced search engines and full-fledged portal sites, users will most likely want small portals providing access to the services they really are interested in, no matter if it comes to business or pleasure. This will lead the way to new opportunities for companies that either understand the customers needs very well, or can personalise such portal sites to meet the demands of each and every customer.

So, what kind of services do users want? Of course, most of the services we are used to today can be of interest in the wireless community as well. As indicated above, the key to successfully launching these services is "utility". If the utility is not high enough it is quite unlikely that the services will be widely used. However, one must remember that the utility of, for example, a game might be very high in certain situations. Do also keep in mind that simple-to-use services often are needed to open the door for more advanced ones since the vast majority of the market is not very familiar even with basic mobile value added services today.

Some examples of such services are:

- **Banking**
  Account statements, paying bills, transfer money between accounts...
- **Finance**
  Stock quotes, buy and sell stocks, interest rates, exchange rates...
- **Shopping**
  Buy everyday commodities, books, records...
- **Gambling**
  Lottery, horse-race betting, poker...
- **Ticketing**
  Book and/or buy air tickets, cinema tickets, concert tickets...
- **Weather**
  Weather forecasts, weather on other locations...

In addition, WAP enables a new category of services that we do not find on the Internet today - telephony services. These services will not only bring utility to the user, which she presumably is willing to pay for, they will also increase the operator's revenues due to increased voice traffic if designed correctly. Examples are:

- **Call management**
  A wide variety of services including:
  - *Incoming call selection*
    Allow the user to choose how an incoming call should be handled. Options could be: answer, reject, forward to assistant, forward to voice mail, etc
  - *Multiparty*
    Provide a comprehensive user interface to multiparty call handling
  - *Call waiting*
    Handle waiting calls with an attractive user interface
  - *Forwarding rules*
    Set and view forwarding rules
- **Voice mail**
  Provide a menu driven user interface to existing voice mail systems
- **Unified Messaging**
  Handle e-mail, faxes, voice mails, etc in a unified manner
- **Enhanced support of legacy SMS services**
  Allows seamless migration from existing text based services into WAP
- **Attractive interface to DTMF services**
  Increase usage of existing DTMF services by providing a better user interface
- **Advanced phonebook management**
  Allows the user to update her phonebook, for example, download a corporate phonebook or the personal phonebook managed via a common WWW browser on a desktop computer

In addition, operators will be able to drastically reduce their costs for customer care. Today substantial amounts of money are spent on voice call centres, where people get questions about their bill, features of a service, etc answered. Many operators have today successfully launched WWW based customer care services as well, allowing the users to access support data on-line. These services can also be designed to speed up the process at traditional call centres by, for example, make the user fill out a questionnaire to pin down the problem before the customer care operator is contacted. This approach does however not solve the problem entirely since customers do often not have access to the WWW when they are on the move, and hence they need to call the call centre anyway to find help. With a WAP based customer care service the customers would be able to get help whenever they want, without having to spend a substantial amount of time waiting for their call to be answered.

WAP Forum is today conducting work in several areas that will facilitate mobile VAS, such as persistent storage, use of smartcards, provisioning, external interfaces, billing, data synchronisation, user-agent profiles, etc. Two areas that will have direct impact on the services we will see in the future are push and telematics.

Abbreviations

ASP	Active Server Pages
CGI	Common Gateway Interface
DTMF	Dual Tone Multi Frequency
GPRS	General Packet Radio Service
HDML	Handheld Device Markup Language
HDTP	Handheld Device Transport Protocol
HTML	HyperText Markup Language
HTTP	HyperText Transport Protocol
ITTP	Intelligent Terminal Transfer Protocol
MAC	Message Authentication Code
OSI	Open System Interconnection
PDA	Personal Digital Assistant
PDU	Packet Data Unit
RAM	Random Access Memory
ROM	Read Only Memory
SMS	Short Message Service
TCP/IP	Transmission Control Protocol/Internet Protocol
TLS	Transport Layer Security
TTML	Tagged Text Markup Language
UDP	User Datagram Protocol
URL	Uniform Resource Locator
WAE	Wireless Application Environment
WAP	Wireless Application Protocol
VAS	Value Added Service
WCMP	Wireless Control Message Protocol
WDP	Wireless Datagram Protocol
WML	Wireless Markup Language
WSP	Wireless Session Protocol
WTA	Wireless Telephony Application
WTAI	Wireless Telephony Application Interface
WTLS	Wireless Transport Layer Security
WTP	Wireless Transaction Protocol
WWW	World Wide Web
XML	Extensible Markup Language